DOCUMENT RESUME

ED 434 683 IR 057 527

AUTHOR	Van der Merwe, Ina; Van Eeden, Welna; Hartzer, Sandra
TITLE	Remote Electronic Resources and the OPAC: Illustrated by the Unisa Library Experience.
PUB DATE	1998-06-00
NOTE	17p.; In: The Challenge To Be Relevant in the 21st Century: Abstracts and Fulltext Documents of Papers and Demos Given at the [International Association of Technological University Libraries] IATUL Conference (Pretoria, South Africa, June 1-5, 1998), Volume 18; see IR 057 503.
AVAILABLE FROM	For full text: <http://educate.lib.chalmers.se/IATUL/proceedcontents/pretpap/vdmerwe.html>.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Libraries; *Access to Information; Bibliographic Records; *Cataloging; Classification; Foreign Countries; Higher Education; Internet; Library Catalogs; Online Catalogs; Subject Index Terms; World Wide Web
IDENTIFIERS	*Electronic Resources; Links (Indexing); University of South Africa; *URLs; Web Sites

ABSTRACT

This paper describes the Unisa (University of South Africa) Library's experience with cataloging remote electronic resources, including electronic journals, electronic text files, online databases, digital images, Unisa campus Web sites, and mailing list discussions. The first section discusses the decision to add bibliographic references for remote electronic resources to the OPAC (Online Public Access Catalog), policies and criteria supporting this decision, resulting changes in workflow, standards, and the bibliographic description of these resources. Providing access to the intellectual content of remote electronic resources is addressed in the second section, including subject headings, form/genre terminology, classification numbers, and summary notes. The third section covers keeping URLs (Uniform Resource Locators) up-to-date, including the use of a linkchecker, and the development of URNs (Uniform Resource Names), the PURL (Persistent URL), and the DOI (Digital Object Identifier). (MES)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

N. Fjallbrant

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

IATUL

REMOTE ELECTRONIC RESOURCES AND THE OPAC : ILLUSTRATED BY THE UNISA LIBRARY EXPERIENCE

ED 434 683

Van der Merwe, Ina, Van Eeden, Welna, and Hartzer, Sandra

Unisa Library, University of South Africa,
PO Box 392, Pretoria, South Africa, 0003

E-mail: vdmerhcj@alpha.unisa.ac.za

veedeiw@alpha.unisa.ac.za

hartzs@alpha.unisa.ac.za

We would like to share Unisa Library's experience with cataloguing remote electronic resources with you. The Library has faced many changes during the past two years, among them the creation of the GAELIC Consortium, conversion to USMARC and the implementation of the INNOPAC Library system. The cataloguing module only came into full operation in February this year. All these events contribute to the Library being able, for the first time, to make meaningful decisions about the handling of Internet resources.

In my contribution to our combined presentation today, I would like to indicate why we decided to add references to remote electronic resources to our Online Public Access Catalogue, the policies and criteria we thought necessary to support this decision, the changes in workflow that resulted and the way that we deal with the bibliographic descriptions of these resources. My colleagues will discuss the organisation of their intellectual content and keeping the URLs up to date.

Bibliographic access : policies and procedure - Van der Merwe, Ina

Introduction

The Internet has been described as a vast and ever-changing reservoir of information with no central locus of control and librarians are acutely aware of the fact that within this maze of intangible, virtual resources, valuable research materials are being made available almost instantaneously all over the networked world. An increasing number of journals are also appearing in electronic format, though some are still accompanied by the printed version.

In addition, the Web has become accessible to anyone who owns a computer and modem and customers with sufficient time, persistence and resources are not dependent on the library any longer to obtain the information they seek. It has in fact, became quite feasible for the library customer of old to develop a personal customised collection. These demands on the Library require a new role for the librarian, from someone who traditionally acted as a gatekeeper to the collection to that of entering into a partnership with customers. Librarians have realised that, in order to stay relevant and to prevent customers from bypassing the library, ownership and access seem to be the way to go. The Unisa Library decided that the OPAC should also be used as a gateway to metadata repositories on the Internet. This will result in the Library's catalogue no longer being merely an inventory of what the library owns, but also a means of access to selected remote resources.

Policy and procedures

IR057527

BEST COPY AVAILABLE

In order to make this possible, we needed to formulate a policy and to identify criteria for guidelines about providing access for remote electronic resources.

We decided that bibliographic references for the following Internet resources would be included in the OPAC:

- electronic journals;
- electronic text files;
    - full text bibliographic information on the library database network;
    - recommended material for Study Services;
    - reference tools;
    - significant electronic research tools;
    - online manuals for professional use;
- online databases;
- digital images;
- Unisa campus websites;
- mailing list discussions which are refereed and cover topics of scholarly interest.

The criteria that they need to meet is necessary to ensure their relevancy for teaching and research at Unisa. The policy furthermore requires that URLs are to be linked-checked and that no printed copies for accessioning purposes are to be made.

It became apparent that the new project would require input and support at all levels and new procedures regarding workflow issues had to be established. This boiled down to close co-operation between subject librarians, IT specialists in the library, serials librarians, cataloguing staff and staff from our Department of Computer Services.

Once an Internet resource has been identified and selected for cataloguing, a printout of the opening screen of the document together with a clear indication of the URL and information on the method of access, the size of the file and possible restrictions for the user, is sent either to the serials librarian depending on whether it is an electronic journal or not, or to one of the cataloguers who for the time being will be responsible for the creation of the surrogate which will display in both catalogues. The text-based OPAC does not support hypertext links and the Internet address on the bibliographic record can only be viewed by the user. However, the web-based catalogue available on the Internet through the University Homepage provides a gateway to the resource through a hypertext link in the record.

It is probable that electronic resources will be incorporated into regular library processing procedures in future.

Standards

At present there is no international standard for the description of metadata. An experiment on the cataloguing of internet resources was conducted by OCLC in 1992, to test and verify the applicability of the second revised edition of the Anglo American Cataloguing Rules and the use of the USMARC Bibliographic Format. With the addition of an 856 field in the MARC format to accommodate the electronic location and access information, in other words the Internet address, these standards were judged sufficient for the cataloguing of these resources. [1] [2] Although we are aware of other metadata schemes such as the Text Encoding Initiative (TEI) and Dublin Core, we decided to adhere to the same standards that are used for the description of conventional library material. In addition, we make use of the twenty first edition of the Dewey Decimal Classification scheme, the Library of Congress Subject Headings , other authoritative documents and the guidelines set out in the Core Bibliographic Record for Computer Files.

Bibliographic description and special problems

In dealing with the description of remote electronic resources, three elements are needed:

- the bibliographic description;
- the access points, both descriptive and subject related;
- an Internet address or Uniform Resource Locator [2]

This is not as straightforward as it sounds, as Internet resources present special problems of their own, namely:

- **Virtual items**
  Cataloguers are dealing with virtual resources. There are no physical items to be examined, as the description is based on a collection of bytes in a computer's memory.

- **Lack of stability**
  They could furthermore be described as moving targets [4, 5] as these resources lack stability. Cataloguers are constantly confronted with decisions about the updating of bibliographic records to match the current state of a resource which can be changed at any time by the homepage owner. The same applies to the URLs which can become invalid without notice. In such instances, the record could just as well be deleted from the catalogue.

- **Restricted access**
  Although many Internet resources are available free of charge, most commercially published electronic journals incur subscription costs like their print counterparts and this can lead to restricted access for the user. Such information needs to be included in the bibliographic description.

Bibliographic record as displayed in the text-based OPAC

CORP AUTHOR	University of South Africa Library.
TITLE	University of South Africa Library [computer file].
FILE INFO	Computer data.
PUBLISHER	Pretoria : University of South Africa Library, 1997.
E-ACCESS	Link to this electronic resource on the Internet http://www.unisa.ac.za/library/index.html
http://purl.unisa.ac.za/oasis/socsci/19563991	
SUBJECT	University of South Africa. Dept. of Library Services.
NOTE	Title from opening screen.
SYS DETAILS	Mode of access: World-Wide Web.
CONTENTS	About the library--Electronic information resources--Training--What's happening--Information for students--A selection from the Internet.
SUMMARY	Profiles the library at the University of South Africa as one of the largest academic libraries on the African Continent. Provides a brief history of the library; the library's mission statement which stresses customer focus; the size of the various collections; the library services and resources available; highlights its multiskilled personnel and their responsive attitude to technological innovation, and the library's supportive function in the University's role as a distance education institution. Provides links to relevant Internet sites.

	LOCATION	CALL NO	STATUS
1	Internet	Online	AVAILABLE

Figure 1

In the above example, Figure 1, we see the bibliographic record or surrogate for the Unisa Library Homepage as reflected in the text-based OPAC. Note the indication of the broad class of material or **GMD** to which an object belongs, enclosed in square brackets, following the title proper.

All items on the Internet are considered published. [6] In cases where no formal publisher statement is given, this information can be derived from the server portion of the Internet address. [7]

As no physical item is being catalogued, it is not necessary to describe the physical extent of the item in terms of screens and so forth. It is also part of our policy that no paper copies should be made. [8], [9]

Although the rules dictate [10] that description must be done from the opening screen, a note indicating this is also required. Other details such as mode of access, contents and a brief objective summary of the purpose and content of the item should be given.

It is imperative to record the Internet address. In the text-based OPAC, this address acts as a reference to the remote electronic resource and can only be viewed by the user. In the Unisa Library, users access the OPAC through the Web.

Bibliographic record as displayed on the WebPac

Corp author	University of South Africa Library.
Title	University of South Africa Library [computer file].
Publisher	Pretoria : University of South Africa Library, 1997.
File info	Computer data
Click on the following to:	
Link to this electronic resource on the Internet	
Connect to http://purl.unisa.ac.za/oasis/socsci/19563991	
LOCATION CALL NO STATUS Internet Online Available	
Subject	University of South Africa. Dept. of Library Services.
Note	Title from opening screen.
Sys detail	Mode of access: World-Wide Web.
Contents	About the library--Electronic information resources--Training--What's happening--Information for students--A selection from the Internet
Summary	Profiles the library at the University of South Africa as one of the largest academic libraries on the African Continent. Provides a brief history of the library; the library's mission statement which stresses client focus; the size of the various collections; the library services and resources available; highlights its

BEST COPY AVAILABLE

Summary	and resources available; highlights its multiskilled personnel and their responsive attitude to technological innovation, and the library's supportive function in the University's role as a distance education institution. Provides links to relevant Internet sites.

Figure 2

In the above example, Figure 2, exactly the same information is displayed, but on the web-based OPAC which is available on the Internet through the University Home Page. The only difference is that the Internet address is now displayed as a hypertext link which acts as a gateway to the online resource. By clicking on this link, direct access is gained to the particular resource. Access to electronic resources can furthermore be enhanced by the inclusion of subject headings, classification numbers and the summary in the note field, and these will be dealt with by Welna van Eeden.

Providing access to the intellectual content of remote electronic resources - Van Eeden, Welna

Through the Internet, scholars and researchers have access to a vast range of dynamic information resources, many of which are not available in any other form. However, automated subject searching on the Internet via search engines does not always guarantee relevance or quality. With the development of Web-interfaced OPACs, which enable users to access remote electronic resources from their library OPACs, cataloguers are challenged to reconsider how the intellectual content of these resources can best be represented and organised. Integrated subject access should be provided through the library OPAC to all formats of information resources. This should include subject access to those parts of the Internet which libraries choose to access via their Webcats. [11]

Subject cataloguing standards and practices have much to contribute to the subject analysis of the information universe found on the Internet. Librarians have been organising the world of knowledge for a long time. Tools such as controlled vocabularies, classification schemes and automated search systems have been developed. These tools could be used, perhaps with minor additions, for the subject cataloguing of Internet resources to ensure proper collocation or arrangement by subject. We should not ignore a century of tradition in subject cataloguing, but continue to build upon that knowledge. [12] Furthermore, the dependence of the library community on shared cataloguing makes it critical that cataloguers abide by the same standards.

The intellectual content of remote electronic resources can be analysed and made accessible to users in Online Library Catalogues:

- by assigning subject headings
- by assigning classification numbers
- by providing a summary in the note fields.

Subject Headings

Why should subject headings be assigned to remote electronic resources?

Although most Web search engines access enormous numbers of hits, research shows that most searchers rarely look at more than two screens of information. Because of this, search engines such as Yahoo are beginning to offer categories of subject searches that can be chosen before making a query. [13]

It is easy to criticize the aimless "surfing" on Web-based search engines without proposing a valid alternative. What are libraries going to do with those parts of the Internet which they want in their Webcats? The alternative could be the application of the basic principles and structures of subject cataloguing to enable a purposeful search and retrieval system which produces results high in relevance

and recall. [14] Furthermore, standardised subject cataloguing will ensure the integration of information resources.

A library may choose the system of subject access it desires - the most widely used being the Library of Congress Subject Headings. The Unisa Library uses Library of Congress Subject Headings as standard for subject cataloguing. This standard will be used to provide subject access to remote electronic resources as well.

Why should a library choose to use Library of Congress Subject Headings?

I would like to answer this question by using the Infomine library as an example. *Infomine*, a virtual library developed by the Library of the University of California, contains close to five thousand records describing and providing access to academically useful Internet resources. Their reasons for chosing Library of Congress Subject Headings as subject vocabulary as follows:

- LCSH uses a standardised descriptive language familiar to librarians in all disciplines. This familiarity can be said to exist for many academic library users as well.
- LCSH is a vocabulary that provides a common thread or a set of controlled access points across the disciplines. The schedule has many complex links between related subjects.
- LCSH is useful for both general and high-level subject description. [15]

LCSH have also been applied successfully in:

- *InterCat*, a searchable catalogue of Internet resources. This catalogue, created by the OCLC Internet Cataloging project, comprises nearly 1,000 bibliographic records of selected Internet resources. [16]
- NetFirst, another initiative of OCLC to provide access to resources on the Internet. The initial release containes 40, 000 records and provides coverage of World Wide web sites, listservs, Usenet groups and anonymous FTP sites. [17]

Assigning Library of Congress Subject Headings to remote electronic resources

On the WWW, everything is represented to the user as a hyper-text object. Hyper-text links in sources make the original object a part of a larger whole. In distinct contrast to the static subject content of a conventional book, electronic resources are dynamic and interactive. [18]

The ALA pre-conference, held in June 1997, with the title: "Demystifying subject cataloging of electronic resources", provided ***General guidelines*** for the subject cataloguing of electronic resources. These guidelines included the following:

- Treat electronic resources as you would printed materials and other formats. It is recommended that libraries treat computer files just like any other form of material in terms of the number and kinds of subject headings assigned.
- Names, chronological elements and form should be part of the subject terminology.
- Subdivisions should be used to make a subject heading more specific, However, a subdivision should not be automatically assigned to indicate that the item is an electronic resource. Form should only be brought out where applicable, e.g. for screen savers. At present there is no form subdivision for ?electronic journals? available in LCSH. This is an area in the cataloguing of remote access serials under discussion. Appropriate subject headings should be subdivided only by the free-floating subdivision - *Periodicals* and should not be further subdivided merely to indicate that the serial is in electronic form.

Many form subdivisions are the same for electronic resources and for printed materials, e.g. sources that have the structure of, and present themselves as reference-type works such as directories, bibliographies

and catalogues. [19]

Form/genre terminology in library OPACS?

When subject headings are assigned to Internet resources, form terminology plays an prominent role. The Library of Congress decided that form/genre access deserves to stand on its own. They are now planning the implementation of a new sub-field for form subdivisions as well as a new field for form/genre headings in authority and bibliographic records. In OPACs, this new field can be indexed separately as a form/genre index alongside author/title/subject access. This will improve OPACs in terms of indexing and display. E.g., it should be possible to distinguish between *Electronic encyclopedias* issued in electronic form (form/genre heading) and works about *Electronic encyclopedias*. [20] Examples of topical headings which the Library of Congress is considering as potential candidates for form headings are Web sites, Electronic discussion groups, Computer fonts and Screen savers (Computer programs)

Classification Numbers

Apart from subject headings, classification numbers can also be used to denote the subject contents of remote electronic resources. If classification numbers are assigned, and the OPAC allows users to browse and search on classification data, users will be able to find electronic resources in the same way as other materials. However, many OPACs provide only call number browsing.

Classification experts and librarians have long recognised the potential of organised, structured library classification schemes to improving subject access to information:

- to improve precision or recall
- to enable browsing
- to serve as a mechanism for switching between languages. [21]

Many libraries do not assign classification numbers to remote electronic resources as they cannot be shelved physically. Michael Gorman, Dean of Library Services at the California State University, believes that, for electronic resources, ?...we should divorce the question of shelf arrangement from that of classification as a retrieval device... Since classification will be used for subject retrieval rather than for identifying a particular book, there is no reason why, when warranted, the cataloguer should not assign more than one classification number to one document... We should explore the power of online systems to coordinate classification numbers, the index to classification schemes, and verbal subject headings to create a subject searching capability that is beyond anything that we have now.? [22] The principle of assigning more than one classification number is applied by the *NetFirst* database.

The OASIS WebPac of the Unisa Library allows call number searches. This index is built from the USMARC 092 tag (call number) which reflects our holdings and not from the USMARC 082 tag (Dewey classification number) We have, however, decided to add the 082 field to records for remote electronic resources (excluding electronic serials). This was decided because of future developments which may take place in organising remote electronic resources according to classification numbers. As the DDC manual gives guidelines for the choice among numbers, we decided to adhere to these principles and only assign one Dewey number per record.

Electronic versions of the DDC and LCC have made it possible to realise the potential of library classification to improve subject retrieval. However, much of the renewed interest in classification schemes as an organising and retrieval device for information resources has been sparked by the growth in usage of the Internet and World-Wide Web. The hierarchical structures of classification schemes support topic browsing. Captions and relative index terms in translation databases could be used to provide a multilingual subject browser to a database of Internet-accessible resources that have been assigned classification numbers. [23]

Several noncommercial World-Wide Web sites are using DDC and LCC to provide subject access to Web-accessible documents, e.g.

- *CyberDewey: a guide to Internet resources*, organised using Dewey Decimal Classification codes. [24]
- *CyberStacks*, a collection of World-Wide Web and other Internet resources that apply the Library of Congress classification scheme to facilitate identification and use of selected sources. [25]

Online classification data can form an important bridge between library methods of organising materials and Internet-based techniques for accessing electronic collections. [26] OCLC has initiated the *Scorpion* research project to address the challenge of applying classification schemes and subject headings in a cost effective way to electronic resources. The idea behind *Scorpion* is that resources to be subject-catalogued can be treated as queries against a special Dewey database that returns a ranked list of potential subjects. [27]

Summary

The summary field is unique to computer file records and can be used to enhance the subject retrieval. One of the weaknesses of most search engines is that relationships and relevance can often not be analysed without actually examining each item. A summary note on the bibliographic record will enable OPAC users to examine the relevance of individual records and then choose among them immediately prior to accessing - thus saving considerable time.

Cataloguers can coordinate the controlled vocabulary in the subject headings field with the natural language in the summary note, particularly for systems that allow for keyword searching of the summary field.
A summary note should provide a brief, objective summary of the purpose and content of the resource. The summary should include such information as:

- the scope of the work
- the nature and form of the contents
- levels of user activity
- age level, degree of difficulty, and/or intended audience. [28]

The catalogue of the future may well provide a summary of the world's information resources. [29]

Cataloguers have long applied their principles and standards of subject cataloguing to materials of various formats. We should be able to apply these principles to the subject cataloguing of remote electronic resources with equal efficiency, to facilitate access to high quality, well selected and annotated sources.

Keeping URLs up to date - Hartzer, Sandra

Introduction

The point-and-click feature of World Wide Web has made Internet browsing so easy that information can be as close as the click of a mouse, but all too often the click leads to a dead end. [30]

URLs often do not work because Internet resources move, names or method of access change and hardware is reconfigured, among other reasons, leaving users stranded with the all too familiar error message 404 of *Document Not Found*. When a URL fails, all instances of that URL (for example, links in a Web document or a bibliographic record) become invalid. [31]

The volatility of Internet resources is an inconvenience at best. For librarians, it is a serious problem which compromises their service to patrons and imposes an unacceptably large burden on catalogue maintenance. [32]

Cataloguers are responsible for creating bibliographic records for Internet resources and with that comes

the responsibility of continually maintaining the accuracy of the location of these resources in library catalogues. How can this best be done? [33]

Solutions to the problem

There are a several possible solutions to the problem:
making use of

- a linkchecker
- URN or PURL
- DOI

Linkchecker

Running a linkchecker against a large library catalogue database is not easy. It is not possible to run the linkchecker against the database itself, which means that all the 856 MARC fields containing URLs have to be extracted, copied into a file and then have the linkchecker run through all the URLs. This method involves a lot of manual work from the cataloguer.

URN

Another solution is the development of Uniform Resource Names, or URNs. The process of defining URNs has been underway in the Internet Engineering Task Force (IETF) for some time where OCLC is an active participant and supporter of this process. [34]

What is the relationship between PURL, URN and URL?

PURLs are a direct result of OCLC's work in the Uniform Resource Name (URN) standards and library cataloguing communities. PURLs satisfy most of the requirements of URNs in the technology that is deployed today. This technology can be applied to the task of maintaining catalogues of Internet resources, and can be smoothly merged into the URN architecture once it is implemented. The assignment of PURLs is an intermediate step towards the time when URNs are an integral part of the Internet information architecture. [35]

Unfortunately the standardisation process is slow. Putting all the pieces in place will require consensus in:

- the IETF,
- the community of Web browser developers and implementors,
- implementation of a new code by the community of network system managers who administer the Domain Name System (DNS) for the Internet. [36]

The concerns and needs of the library community may not be fully appreciated or adequately addressed by these groups quickly enough. Libraries can and should therefore provide leadership in the solution of these problems.

To aid in the development and acceptance of URN technology, OCLC has developed a naming and resolution service for general Internet resources. The names, which can be thought of as Persistent URLs (PURLs), can be used in documents, Web pages and in cataloguing systems. [37]

PURL

What is a PURL?

PURL is the acronym for "Persistent Uniform Resource Locator."

"Functionally, a PURL is a URL. However, instead of pointing directly to the location of an Internet

resource, a PURL points to an intermediate resolution service. The PURL Resolution Service associates the PURL with the actual URL and returns that URL to the client. The client can then complete the URL transaction in the normal fashion." [38] In the Web environment, this is a standard HyperText Transfer Protocol (HTTP) redirect. PURLs have been assigned to records catalogued in the Internet Cataloging Project ($g in 856 USMARC tag)

The OCLC PURL Service has been running since the beginning of January, 1996.

Although a PURL Service is being run and maintained at OCLC, the PURL source code is freely distributed on the Internet to aid in the rapid distribution of this technology and since the introduction of PURL, a number of institutions have expressed an interest in running their own PURL servers. [39]

What does a PURL do?

A user selects (clicks on) a PURL on a Web page, in a document or a bibliographic record, the PURL then resolves the associated URL, which the browser then in turn uses to access the resource. [40]

A PURL can be used with the confidence that it will persist over time. The links will remain valid even though the associated URLs might change. This does not mean that a PURL magically changes its associated URL when the referenced resource moves -- the maintainers of the PURL make this happen. [41]

What makes a PURL Persistent?

PURLs never change. This means that a PURL can last longer than any particular URL that may be associated with it. While PURLs allow you to associate different URLs with them, the PURL itself never changes. In other words , you can change what a PURL resolves to, but you cannot change the PURL. [42]

Because a PURL persists indefinitely, all instances of such a PURL will remain valid. If the associated URL of a PURL becomes outdated, resolution of the PURL may fail. Of course, someone has to operate the PURL resolvers that provide this persistence. This is where OCLC plays a big role. As part of OCLC's commitment to facilitating access to the world's information, they are encouraging the use of PURLs. [43]

It is expected that organisations with similar commitments to providing long term access to information will want to run PURL Servers as well, for example:

- government agencies
- publishers
- libraries
- universities

It is therefore important to understand that persistence is a function of the organisation, and not the technology. [44]

What does a PURL look like?

A PURL looks just like a URL because it is a URL. A PURL is made up of three parts:

- the protocol
  The protocol is used to access the PURL resolver. It is important to note that this protocol may differ from the one used to access the resource associated with the PURL, e.g. it is possible to have a PURL (http protocol) for a gopher site.

- the resolver address

This is the IP address or domain name of the PURL resolver. This part of the PURL is resolved by the Domain Name Server (DNS).

- name
  A user-assigned name. This name may differ from the name of the resource in the associated URL.

Here are a few examples of PURLs.

```
http://purl.unisa.ac.za/oasis/natsci
----   ------------ ---------
  /          |          \
protocol  resolver address   name

http://purl.unisa.ac.za/oasis/socsci
----   ------------ --------------------
  /          |              . \
protocol  resolver address       name [45]
```

What is a PURL resolver?

"A PURL resolver is a service, available via standard HTTP 1.0 protocols, that facilitates the creation, maintenance, and resolution of PURLs." [46]

What are PURL domains?

Domains are subdivisions of the name space on a PURL resolver. They are very much like directories in a file system. There are two varieties of domains, namely top-level domains and subdomains.

- a top-level domain is the top-level of the name space on a PURL resolver.
- Subdomains exist within top-level domains or other subdomains. Subdomains can be created in any domain.

For example, the PURL http://purl.unisa.ac.za/oasis/socsci has 2 domains: oasis and socsci; oasis is a top-level domain, socsci is a subdomain of oasis [47]

Which resources should have PURLs?

Users should assign a PURL to any resource for which reliable access over time is desired. For example:

- a home page
- an electronic journal or book
- an individual article or a paper are also good candidates for a PURL.

Resources that do not justify PURLs include:

- sections within a document
- charts or graphics that would not make sense outside the context of their containing document
- temporary resources. [48]

Creation and maintenance of PURLs

To create a PURL, the user should access a PURL Resolver via a Web browser and follow the resolver's instructions for creating a PURL. PURL Resolvers provide a form to fill in in order to create a PURL.

The question arises: are PURLs updated automatically when their associated URL changes? The answer

 BEST COPY AVAILABLE

to this question is no, it is the responsibility of a PURL's owner and its maintainers, the cataloguers, to update the PURL when the associated URL changes. Each maintainer is responsible for certain PURLs.

To update or maintain a PURL the PURL resolver's maintenance forms are used to make the appropriate changes to the desired PURL [49]

PURL Source Code

The software, which includes the source code, can be downloaded from the OCLC PURL site. Full instructions are given on how to install the software and also the hardware needed. The later versions of the PURL software also include a validator.

A mailing list is also available. For further PURL information, the PURL FAQ at http://purl.oclc.org/OCLC/PURL/FAQ is very useful. [50]

DOI (Digital Object Identifier)

DOI is an identification system for electronic resources, a universal accession number that never changes. DOI was developed by the Association of American Publishers in collaboration with the non-profit Corporation for National Research Initiatives. Several publishers are taking part in a pilot project which started in July 1997. John Wiley and Sons have already used DOIs in 63,000 items. [51]

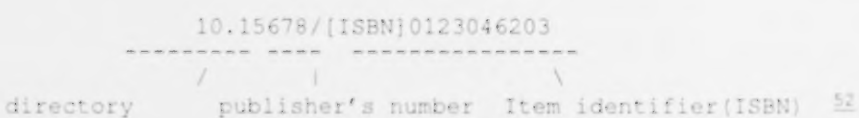
The structure of DOI

The DOI system is made up of a prefix and a suffix

- Prefix
  The prefix indicates the DOI directory, e.g. 10, which indicates a US directory, and the publisher's number. This number is assigned by the Corporation for National Research Initiatives, e.g. 1006 for Academic Press.

- Suffix
  The suffix is the item identifier that the publisher assigns to a particular item. Some publishers use the ISBN as the suffix, with a code indicating that an ISBN is used. Other publishers use their own internal numbering system.

Structure of a DOI



[52]

Benefits of DOI

- It is persistent over time
- It will function in conjunction with the proposed URN system
- Although it is still in its initial stages of evaluation, it is proving to be working well. [53]

Information about getting started with DOI can be found at : http://www.doi.org/started.html

We as librarians owe it to our users to make sure that what we have or what we provide access to in our library catalogue is retrievable. Whether we do it with PURL, URN, linkchecker or DOI, time will tell.

Conclusion

BEST COPY AVAILABLE

We librarians are living in exciting times - what with the multitude of technological advances taking place around us and the developments in information distribution and access. We should grasp all the opportunities these developments provide to enhance our service and to meet our users' increasingly sophisticated information needs. None the less, we will only have succeeded once the user recognises the additional benefits of this service.

References

1. GUENTHER, Rebecca S. The challenges of electronic texts in the library : bibliographic control and access. In Clinic on Library Applications of Data Processing (31st : 1994 : University of Illinois at Urbana-Champaign) Literary texts in an electronic age. University of Illinois at Urbana-Champaign, 1994, pp. 149-172.

2. HEERY, Rachel. Review of metadata formats (IAFA/Whois++, MARC, Text Encoding Initiative, Dublin Core, Uniform Resource Characteristics) Program, 30, 1996: pp. 345-373.

3. FATTIG, Karl. BLOCK, Rick J. Cataloging internet resources : policy, procedure and practice. Technicalities, 17, 1997: pp. 1, 3-6.

4. MOIR, Susanne, WELLS, Andrew. Descriptive cataloguing and the Internet : recent research. Cataloguing Australia, 22(1/2) 1996: pp. 8-16.

5. OLSON, Nancy B. Cataloguing Internet resources [online]. 2nd ed. Dublin, Ohio : OCLC, 1995. [cited 20 April, 1998]. Available from the World Wide Web: http://www.purl.org/oclc/cataloging-internet.

6. OLSON, ref 5.

7. FATTIG, ref 3, p. 5.

8. OLSON, ref 5.

9. FATTIG, ref 3, p. 5.

10. ANGLO-AMERICAN CATALOGUING RULES. Prepared under the direction of the Joint Steering Committee for Revision of AACR. Edited by Michael Gorman and Paul W. Winkler. Revised ed. Chicago : American Library Association, 1988.

11. SHA, Vianne T. Cataloguing Internet resources: the library approach. The Electronic Library, 13(5) 1995: pp. 467-476.

12. BUTTERFIELD, Kevin L. Cataloger's and the creation of metadata systems: a collaborative vision at the University of Michigan. In Proceedings of the OCLC Internet cataloging project colloquium, San Antonio, Texas, January 19, 1996 [online]. Dublin (Ohio) : OCLC, 1996. [24 April 1998]. Available from World Wide Web: http://www.oclc.org/oclc/man/colloq/butter.htm.

13. Mary Micco discusses LC's role in Internet subject retrieval. In LC Cataloging Newsline: Online Newsletter of the Cataloging Directorate, Library of Congress 5(12) December 1997 [online]. Washington, DC : Library of Congress, 1997.[cited 24 April 1998]. Available from World Wide Web: http://www.loc.gov/catdir/lccn/lccn0512.html.

14. GORMAN, Michael. What is the future of cataloguing and cataloguers. In 63rd IFLA General Conference, August 31-September 5, 1997: conference programme and proceedings [online] . The Hague, Netherlands : IFLA, 1997. [cited 24 April 199]. Available from World Wide Web: http//www.nlc-bnc.ca/ifla/IV/ifla63/63gorm.htm.

15. MITCHELL, Steve. Library of Congress Subject Headings as subject terminology in a virtual

library: the INFOMINE example. In Untangling the Web: proceedings of the Conference sponsored by the Librarians Association of the University of California, Santa Barbara and friends of the UCSB Library, April 26, 1996, University Center, University of California, Santa Barbara [online]. Santa Barbara : University of California, 1996. [cited 24 April 1998]. Available from World Wide Web: http://www.library.ucsb.edu/untangle/smitch.html.

16. Internet cataloging project database now available--July 21, 1995. In OCLC news releases. Collections and Technical Services news releases [online]. Dublin, Ohio : OCLC, 1995. [cited 24 April 1998]. Available from World Wide Web: http://www.oclc.org/oclc/press/950721a.htm.

17. Netfirst database introduced. In OCLC news releases. Reference Services news releases [online]. Dublin (Ohio) : OCLC, 1996. [cited 23 April 1998]. Available from World Wide Web: http://www.oclc.org/oclc/press/960219b.htm.

18. BARKER, Philip. Living books and dynamic electronic libraries. The Electronic Library, 14(6) 1996: pp. 491-500.

19. RESER, David. Assigning LCSH to electronic resources. Demystifying subject cataloging of electronic resources: ALCTS preconference, San Francisco, California, 26-27 June 1998. Unpublished.

20. YEE, Tom. Report to ALA ALCTS CCS Subject Analysis Committee, Subcommmittee on form headings/subdivisions implementation, June 28, 1997. In ALCTS/CCS/SAC Subcommittee on form headings/subdivision implementation, 1996 - LC report (June 1997) [online]. Pittsburgh : University of Pittsburgh, 1997. [cited 24 April 1998]. Available from World Wide Web: http://www.pitt.edu/~agtaylor/ala/lcre972.htm.

21. VIZINE-GOETZ, Diane. Using library classification schemes for Internet resources. In Proceedings of the OCLC Internet Cataloging Project Colloquium, San Antonio, Texas, January 19, 1996 [online]. Dublin, Ohio : OCLC, 1996. [cited 24 April 1998]. Available from World Wide Web: http:///www.oclc.org/oclc/man/colloq/v-g.htm.

22. GORMAN, ref. 14.

23. VIZINE-GOETZ, ref. 21.

24. CyberDewey [online]. Dublin, Ohio : OCLC, 1989. [cited 4 April 1998]. Available from World Wide Web: http://ivory.lm.com/~mundie/DDHC/DDH.html - CyberDewey.

25. CyberStacks(sm) [online]. Ames, Iowa : Iowa State University, 1998. [cited 24 April 1998]. Available from World Wide Web: http://www.public.iastate.edu/~CYBERSTACKS.

26. VIZINE-GOETZ, ref. 21.

27. SHAFER, Keith E. Automatic Subject Assignment via the Scorpion System [online]. Dublin, Ohio : OCLC, 1996. [cited 24 April 1998]. Available from World Wide Web: http://www.oclc.org/oclc/research/publications/review96/scorpion.htm.

28. RESER, ref. 19.

29. TYCKOSON, David A. The twenty-first century limited: designing catalogs for the next century. Cataloging and classification quarterly, 13(3-4 1991, pp. 3-27.

30. SHAFER, Keith, et al. Introduction to Persistent Uniform Resource Locators [online]. Dublin, Ohio : OCLC, 1996. [cited 27 April 1998]. Available from World Wide Web: http://purl.oclc.org/OCLC/PURL/INET96.

31. SHAFER, ref 30.

32. WEIBEL, Stuart, Jul, Erik, and Shafer, Keith. PURLs: Persistent Uniform Resource Locators [online]. Dublin, Ohio : OCLC, 1995. [cited 28 April 1998]. Available from World Wide Web: http://purl.oclc.org/OCLC/PURL/SUMMARY.

33. SHA, Vianne T. Cataloguing Internet resources: the library approach. The Electronic Library, 13(5) 1995, p. 472.

34. WEIBEL, ref 32.

35. WEIBEL, ref 32.

36. WEIBEL, ref 32.

37. WEIBEL, ref 32.

38. SHAFER, ref 30.

39. 31. SHAFER, ref 30.

40. PURL Frequently Asked Questions [online]. Dublin, Ohio : OCLC, 1996. [cited 28 April 1998]. Available from World Wide Web: .

41. SHAFER, ref 30.

42. PURL Frequently Asked Questions, ref 40.

43. SHAFER, ref 30.

44. SHAFER, ref 30.

45. SHAFER, ref 30.

46. PURL Frequently Asked Questions, ref 40.

47. PURL Frequently Asked Questions, ref 40.

48. PURL Frequently Asked Questions, ref 40.

49. PURL Frequently Asked Questions, ref 40.

50. PURL Frequently Asked Questions, ref 40.

51. BERINSTEIN, Paula DOI: a New Identifier for Digital Content [online]. [S.l. : sn , 199- [cited 28 April 1998]. Available from World Wide Web: http://www.infotoday.com/searcher/jan/story4.htm.

52. DOI Prototype FAQ [online]. [S.l. : Academic Press, 1997. [cited 28 April 1998]. Available from World Wide Web: http://www.apnet.com/www/doi/faq.htm.

53. DOI Prototype FAQ, ref 52.

List of acronyms

ALA - American Library Association
DDC - Dewey Decimal Classification
DOI - Digital Object Identifier

GAELIC - Gauteng and Environs Library Consortium
LCC - Library of Congress Classification
LCSH - Library of Congress Subject Headings
OCLC - Online Computer Library Center
OPAC - Online public access catalogue
PURL - Persistent Uniform Resource Locator

Biographical details

Ina van der Merwe and **Welna van Eeden**, both from the Standards Team of Technical Services in the Unisa Library, are responsible for the co-ordination of activities to ensure the adherence to the uniform application of standards such as the USMARC Format, DDC and LCSH. They are also involved in the Cataloguing Task Group of the GAELIC Consortium.

Sandra Hartzer is the Pagemaster for the Unisa Library Home page and Intranet and is the co-ordinator of Internet Training for the Unisa Library.

Back to contents

Last edited by J.F, 30th June, 1998.